

# Proper Motion Catalog From SDSS $\cap$ USNO-B

Andrew Gould and Juna A. Kollmeier

*Department of Astronomy, The Ohio State University, 140 W. 18th Ave., Columbus, OH  
43210*

`gould,jak@astronomy.ohio-state.edu`

## ABSTRACT

We present a catalog of 345,000 stars with proper motions  $\mu \geq 20 \text{ mas yr}^{-1}$  and magnitudes  $r' \leq 20$ , drawn from the intersection of the SDSS Data Release One and USNO-B. We first use SDSS quasars to correct the proper motions in each of the two source catalogs, finding that these corrections for systematics are comparable to the statistical errors. The combined proper motions have errors of  $\sigma_\mu = 4.0 \text{ mas yr}^{-1}$ , significantly better than any previous catalog at these magnitudes. While cross-correlating SDSS and USNO-B removes the vast majority of the very large number of spurious proper-motion stars in each, some contamination remains, especially at faint magnitudes,  $r' \gtrsim 19.5$ , and high proper motions. We present a diagrammatic method to estimate the level of contamination in selected subsets of the catalog, which should be useful for choosing appropriate selection criteria for various applications.

*Subject headings:* astrometry – catalogs – stars: fundamental parameters – stars: subdwarfs – late-type – white dwarfs

## 1. Introduction

Combining stellar proper motions with photometry has long been an efficient method to select otherwise elusive populations of stars. Proper motion surveys can be broadly divided into two classes: those that cover a large fraction of the sky and those that cover small areas but have faint limiting magnitudes.

All-sky surveys generally trade magnitude limits for proper motion precision or threshold. For example, the combined Hipparcos (ESA 1997) and Tycho 2 (Høg et al. 2000) catalogs provide proper motions accurate to  $\sigma_\mu \lesssim 1 \text{ mas yr}^{-1}$  for essentially all stars with  $V \lesssim 12$ . The nearly completed USNO CCD Astrograph Catalogs (UCAC1,2) (Zacharias et al. 2000,

2001) have  $\sigma_\mu$  between  $1 \text{ mas yr}^{-1}$  and  $5 \text{ mas yr}^{-1}$  depending on magnitude for  $V \lesssim 16$ . The New Luyten Two Tenths (NLTT) Catalog (Luyten 1979, 1980; Luyten & Hughes 1980) is reasonably complete to  $V \sim 19$  over most of the sky, but at the price of being restricted to high proper-motion stars ( $\mu > 180 \text{ mas yr}^{-1}$ ). Moreover, at  $\sigma_\mu \sim 20 \text{ mas yr}^{-1}$  (Salim & Gould 2003), its accuracy is substantially worse than catalogs at brighter magnitudes. However, the revised NLTT (rNLTT) improves this precision to  $\sigma_\mu = 5.5 \text{ mas yr}^{-1}$  for 44% of the sky (Gould & Salim 2003; Salim & Gould 2003). Most recently, Monet et al. (2003) have analyzed 4 decades of photographic plates to produce the monumental USNO-B, an all-sky proper-motion catalog with typical  $\sigma_\mu \sim 7 \text{ mas yr}^{-1}$  for all stars (no proper-motion threshold) down to  $V \sim 21$ . However, while this catalog is 90% complete for high proper motion stars ( $\mu \gtrsim 180 \text{ mas yr}^{-1}$ ), about 99% of its entries in this regime are spurious (Gould 2003c), so it cannot be directly accessed as a source of high proper motion stars.

Due to the relatively bright magnitude limits of these all sky surveys (other than USNO-B), other groups have been motivated to do smaller-area, deeper studies, focusing instead on a particular proper-motion or magnitude limit as opposed to all-sky coverage. The SuperCOSMOS catalog (Hambly et al. 2001a,b), covering a 5000 square degree patch of the southern Galactic cap, obtains proper motion measurements with  $\sigma_\mu \sim 10 \text{ mas yr}^{-1}$  at  $R \sim 18$ , and  $\sigma_\mu \sim 50 \text{ mas yr}^{-1}$  at  $R \sim 21$ . Fainter still, is the Calán-ESO proper motion catalog (Ruiz et al. 2001), which contains 542 stars with  $\mu > 200 \text{ mas yr}^{-1}$  to  $R \sim 19.5$  in a 350 square degree field in the south. Also in the south is the Wroblewski & Costa (2001) survey of 147 stars with  $\mu > 150 \text{ mas yr}^{-1}$  ( $\sigma_\mu \sim 6 - 22 \text{ mas yr}^{-1}$ ) over a 25 square degree field down to  $B \leq 18.5$ . Using the Digitized Sky Survey (DSS), Lépine, Shara, & Rich (2003) find proper motions for stars between  $500 \leq \mu \leq 2000 \text{ mas yr}^{-1}$  accurate to  $10 \text{ mas yr}^{-1}$  with  $8 \leq R \leq 20$  covering over 98% of the northern sky.

Finally, several proper motion catalogs have been produced as byproducts of searches for halo dark matter and therefore have wholly different selection functions than the surveys mentioned above. The EROS 2 HPM catalog (Goldman et al. 1999) selects stars from 413 square degrees at high Galactic latitude to  $V \leq 21.5$  and  $I \leq 20.5$  with an accuracy of  $25 \text{ mas yr}^{-1}$ . The MACHO survey provided a catalog of 154 high proper-motion ( $\mu > 5 \text{ mas yr}^{-1}$ ) stars from a total of 50 square degrees in fields towards the Galactic bulge and the LMC with an accuracy of  $\sim 0.8 \text{ mas yr}^{-1}$ . The Optical Gravitational Lensing Experiment (OGLE II) catalog (Sumi et al. 2003) yielded proper motions for 5,078,188 stars with  $\mu \leq 500 \text{ mas yr}^{-1}$  and accuracy  $\sigma_\mu \sim 0.8 - 3.5 \text{ mas yr}^{-1}$  from a total imaging area of 11 square degrees towards the Galactic bulge.

These catalogs have fueled many different veins of study. Searches for a variety of distinct stellar populations including nearby stars (Reid, Liebert, & Schmidt 2001; Jahreiss

et al. 2001; Scholz, Meusinger, & Jahreiss 2001; Gizis & Reid 1997; Henry, et al. 1997), subdwarfs (Gizis & Reid 1997; Ryan 1992; Digby et al. 2003), white dwarfs (Reid, Sahu, & Hawley 2001; Schmidt et al. 1999; Liebert et al. 1979; Jones 1972; Luyten 1970, 1977), brown and L dwarfs (Goldman et al. 1999), halo stars (Gould 2003a), and wide binaries (Chaname & Gould 2004) are all most effectively carried out using proper motion catalogs such as those described above. Proper motion catalogs have also been useful as a way to select candidates for future microlensing events detectable with next generation surveys (Salim & Gould 2000) such as the *Space Interferometry Mission* as well as candidates for planetary transit searches (Gould & Morgan 2003). Searches for dark matter in the halo in the form of white dwarfs (Oppenheimer et al. 2001; Reid, Sahu, & Hawley 2001) and MACHOs (Yoo et al. 2004) have also benefited from proper motion catalogs. Finally, the structure of the stellar halo including the velocity ellipsoid parameters (Gould 2003a) and the granularity of the stellar halo (Gould 2003b) have both been made possible by the large samples of halo stars culled from high proper-motion catalogs.

In this paper, we construct a new proper-motion catalog by combining data from the Data Release One (DR 1) of the Sloan Digital Sky Survey (SDSS, <http://www.sdss.org>) and USNO-B1.0 (Monet et al. 2003). SDSS contains photometry and proper motions for stars over several disjoint areas at high Galactic latitude totaling several thousand square degrees. SDSS assigns proper motions to most of these stars by identifying each with the nearest entry in USNO-A2.0 (Monet 1998), a position catalog based on circa 1950 simultaneous blue and red photographic plates. Many of these proper motions are spurious, either because the SDSS star has no counterpart in USNO-A (and so is identified with an unrelated star) or because it has moved so far that the nearest USNO-A entry is not the true counterpart of this star. Moreover, these proper motions suffer systematic errors due to small but significant zonal biases in USNO-A. USNO-B also contains many spurious high proper-motion entries that arise from inadvertent association of unrelated stars that have moved significantly between plate epochs, and it also suffers from zonal errors.

Here we use the SDSS quasars to remove the zonal errors from each of these two source catalogs. Because the SDSS spectroscopic survey covers only about half the area of the photometric survey, our catalog is restricted to this smaller region in which the quasars are reliably identified. We then combine the two corrected proper-motion measurements to form the single best estimates of the proper motions, which have errors of  $\sigma_\mu \sim 4 \text{ mas yr}^{-1}$ .

The very process of cross-identifying SDSS and USNO-B eliminates most of the spurious entries in each. We develop a graphical technique to estimate the remaining contamination in various subsets of the catalog in order to assist in the formulation of selection procedures appropriate to various applications. Our final catalog contains 345,000 entries with proper

motions  $\mu \geq 20 \text{ mas yr}^{-1}$  and magnitudes  $r' \leq 20$ .

Digby et al. (2003) were the first to carry out a cross-identification of SDSS with an independent proper-motion catalog, namely SuperCOSMOS. They did not attempt to combine the two proper-motion measurements, but rather used the SDSS proper motions as a cross check. They were mostly interested in combining SDSS photometry with their proper motions in order to produce a reduced proper motion diagram from which halo stars could be identified. They achieved a precision of  $\sigma_\mu \sim 8 \text{ mas yr}^{-1}$  and set their catalog threshold at  $\mu \geq 40.5 \text{ mas yr}^{-1}$ .

In § 2 we describe our initial SDSS sample selection. In § 3 we develop our method for correcting SDSS and USNO-B proper motions using SDSS quasars, and we use these quasars to evaluate the error properties of both these catalogs as well as of our final adopted proper motions. In § 4 we describe the construction of the stellar proper-motion catalog, and in § 5 we develop visual techniques for understanding its statistical properties. We also present reduced proper motion diagrams. Finally, in § 6 we describe the electronic catalog.

## 2. SDSS Selection

We select stars and quasars from the SDSS DR1 using the online query tools provided by SDSS. (For technical references, see <http://www.astro.psu.edu/users/dps/sdsstechrefs.html>.) We select all quasars that have matches within  $2''$  to corresponding sources within the USNO-A catalog and that are spectroscopically confirmed quasars. For our star sample, we initially select all objects within DR1 that 1) are classified as stars by SDSS photometrically, 2) have proper motions (measured from USNO-A and SDSS) in the range  $7 \text{ mas yr}^{-1} \leq \mu \leq 420 \text{ mas yr}^{-1}$ , and 3) have magnitudes  $r' < 22$ . We then supplement this with an additional query covering  $\mu > 420 \text{ mas yr}^{-1}$  but restricted to  $r' < 20$ .

## 3. Quasar-Based Calibration

We use the quasars for three related purposes. First, to remove the (position-dependent) astrometric biases in SDSS/USNO-A and USNO-B proper motions. Second, to measure the errors in these proper motions and the covariances among them. Third, as a benchmark of accurately identified objects against which we can compare the distributions of variously selected stellar subsamples in order to judge whether they are significantly contaminated by spurious objects. The quasars have two characteristics that permit them to perform these

three functions. First, they are spectroscopically confirmed by SDSS. Second, they have known proper motions: namely zero.

### 3.1. Individual Catalog Analyses

We begin by correcting the SDSS proper motions by using neighboring quasars to measure the local proper-motion bias. We determine this bias by taking the weighted average of all other quasars,

$$\vec{\mu}_{\text{bias}}^m = \frac{1}{w_{\text{tot}}^m} \sum_{n \neq m} w_n^m \vec{\mu}^n, \quad w_{\text{tot}}^m = \sum_{n \neq m} w_n^m, \quad w_n^m \equiv \exp[-(\vec{\theta}^m - \vec{\theta}^n)^2 / 2\theta_0^2], \quad (1)$$

where  $\vec{\theta}^n$  is the vector position of the  $n$ th quasar and  $\theta_0$  is a smoothing length. We find that putting  $\theta_0 = 0.75$  deg minimizes the residual scatter. In addition to excluding the star that is being corrected from the sum, we also exclude all those with  $R_{\text{USNO}} > 19.5$  or  $R_{\text{USNO}} < 15$ , although this has only a very minor effect. We find that this correction improves the scatter in the quasars by

$$(\sigma_\alpha, \sigma_\delta)_S : (5.2, 5.8) \text{ mas yr}^{-1} \rightarrow (4.7, 4.7) \text{ mas yr}^{-1} \quad (\text{SDSS}), \quad (2)$$

for the sample of 14,740 quasars.

We then apply the same procedure to USNO-B. Of course, we exclude all quasars that are not detected in USNO-B, but we also exclude those that are recorded as having proper-motion errors exceeding  $10 \text{ mas yr}^{-1}$  in either direction, as well as those with zero errors in both components. About 99% of these latter objects are listed as having exactly zero proper motion. Since, as we show below, the bias at the location of these quasars is decidedly non-zero, these entries must be catalog artifacts rather than highly precise measurements. These two cuts reduce the sample to 3543 quasars, with more than 90% of the reduction being the exclusion of the zero-error quasars. Finally, because the quasar sample is now fairly sparse, it is possible that some quasars do not have enough near neighbors to accurately measure their bias. We therefore exclude an additional 175 quasars with  $w_{\text{tot}}^m < 3$ , leaving a sample of 3368. For these,

$$(\sigma_\alpha, \sigma_\delta)_U : (7.9, 7.1) \text{ mas yr}^{-1} \rightarrow (5.1, 5.0) \text{ mas yr}^{-1} \quad (\text{USNO} - \text{B}). \quad (3)$$

Thus, the local bias (and hence the improvement when this bias is removed) is substantially larger for USNO-B than for SDSS/USNO-A. When we measure the bias for the zero-error quasars, it is similar, thus motivating their above-stated exclusion from the analysis.

### 3.2. Joint Catalog Analysis

Since we will eventually measure proper motions for our stellar sample using both SDSS/USNO-A and USNO-B, it is important to understand the joint error properties of all four quantities being measured, which we group together as a vector  $a_i$ ,

$$a_i^m \equiv (\mu_{\alpha,\text{cor},S}^m, \mu_{\alpha,\text{cor},U}^m, \mu_{\delta,\text{cor},S}^m, \mu_{\delta,\text{cor},U}^m), \quad (4)$$

where  $\mu_{\alpha,\text{cor},S}^m \equiv \mu_{\alpha,S}^m - \mu_{\alpha,\text{bias},S}^m$ , and similarly for the  $\delta$  component. We first evaluate the covariances  $c_{ij}$  of the  $a_i$  averaging over the 3368 quasars common to both samples,

$$c_{ij} \equiv \langle a_i a_j \rangle = \frac{1}{3368} \sum_{m=1}^{3368} a_i^m a_j^m. \quad (5)$$

Note that while one would usually write  $c_{ij} = \langle a_i a_j \rangle - \langle a_i \rangle \langle a_j \rangle$ , in the present case the  $\langle a_i \rangle$  are known a priori to vanish. We find errors

$$\sqrt{c_{ii}} = (4.13, 5.15, 4.27, 4.96) \text{ mas yr}^{-1}, \quad (6)$$

and correlation coefficients,

$$\frac{c_{ij}}{\sqrt{c_{ii}c_{jj}}} = \begin{pmatrix} 1.000 & 0.551 & -0.034 & -0.052 \\ 0.551 & 1.000 & -0.065 & -0.080 \\ -0.033 & -0.065 & 1.000 & 0.554 \\ -0.052 & -0.080 & 0.554 & 1.000 \end{pmatrix}. \quad (7)$$

Note that the SDSS scatter is slightly reduced compared to equation (2), indicating that the quasars accessible to USNO-B are somewhat better behaved than the full sample. Since we will also only be considering stars with entries in both catalogs, the numbers quoted in equation (6) are the directly relevant ones.

Equation (7) shows that the SDSS/USNO-A and USNO-B measurements are correlated at the 55% level. Presumably, this is due to the fact that the first Palomar Observatory Sky Survey (POSS I) underlies the first (circa 1950) epoch of both measurements. Also note that the correlations between the  $\alpha$  and  $\delta$  components, while small, are statistically significant because the uncertainties in these quantities are  $(3368)^{-1/2} = 0.017$ . We therefore keep these cross terms in our calculations, which we describe in the appendix. However, while this mathematical analysis is straightforward, it is a bit dense. We therefore present a simplified treatment here that ignores these small terms and so is much more transparent while yielding almost exactly the same results

The internal error is estimated by determining how well the SDSS/USNO-A and USNO-B measurements agree with each other, while the external error is estimated by determining

how well the combined proper-motion measurements agree with reality (i.e., that the proper motions are actually zero).

The internal error is characterized by,

$$\chi_{\text{int},m}^2 = \left[ \frac{\Delta\mu_\alpha^m}{\sigma(\Delta\mu_\alpha)} \right]^2 + \left[ \frac{\Delta\mu_\delta^m}{\sigma(\Delta\mu_\delta)} \right]^2, \quad (8)$$

where  $\Delta\mu_\alpha^m = a_1^m - a_2^m = \mu_{\alpha,\text{cor},S}^m - \mu_{\alpha,\text{cor},U}^m$ , where  $\sigma(\Delta\mu_\alpha)$  is the uncertainty in this quantity, and similarly for the  $\delta$  component. Since the two measurements are correlated, the error in their difference is given by  $[\sigma(\Delta\mu_\alpha)]^2 = c_{11} - 2c_{12} + c_{22}$ . Substituting the values given in equations (6) and (7) yields

$$[\sigma(\Delta\mu_\alpha), \sigma(\Delta\mu_\delta)] = (4.49, 4.40) \text{ mas yr}^{-1}, \quad (9)$$

That is, because the correlation coefficients between the SDSS/USNO-A and USNO-B measurements (eq. [7]) are so large, the errors in the differences of these measurements are only of order the errors in each measurement separately (rather than being multiplied by  $2^{1/2}$ ).

The distribution of  $\chi_{\text{int}}^2$  for the 3368 quasars is shown in the lower panel of Figure 1. The straight line is the prediction for Gaussian statistics, i.e., a  $\chi^2$  of two degrees of freedom, which is an exponential distribution,  $f(\chi^2) = \exp(-\chi^2/2)$ . This line is not a fit: it has no free parameters. The figure shows that the internal errors are Gaussian distributed in the core,  $\chi_{\text{int}}^2 \lesssim 8$ , and then deteriorate to a much flatter distribution in the wings,  $f(\chi^2) \propto \exp(-\chi^2/7)$ . This deterioration is not unexpected and is typical of many types of astronomical observations.

Next we evaluate the external errors. Let  $\tilde{\mu}_\alpha^m$  be the best estimate of  $\mu_\alpha^m$  given the two correlated measurements from SDSS and USNO-B, respectively. Let  $\sigma(\tilde{\mu}_\alpha)$  be the uncertainty in this quantity, and similarly for the  $\delta$  component. Since the quasar proper motion is known to vanish, these “best estimates” are exactly equal to the error in the measurement. Hence, the external errors are characterized by,

$$\chi_{\text{ext},m}^2 = \left[ \frac{\tilde{\mu}_\alpha^m}{\sigma(\tilde{\mu}_\alpha)} \right]^2 + \left[ \frac{\tilde{\mu}_\delta^m}{\sigma(\tilde{\mu}_\delta)} \right]^2. \quad (10)$$

For two correlated measurements  $a_1$  and  $a_2$  of the same quantity, the best combined estimate and its uncertainty are given by,

$$\tilde{a} = \frac{a_1(c_{22} - c_{12}) + a_2(c_{11} - c_{12})}{c_{11} - 2c_{12} + c_{22}} \quad \tilde{\sigma}^2 = \frac{c_{11}c_{22} - c_{12}^2}{c_{11} - 2c_{12} + c_{22}}. \quad (11)$$

These equations can be derived using the formalism presented in the Appendix. However, it is easy enough to verify that they reduce to familiar results when *either*  $c_{12} = 0$  *or*  $c_{11} = c_{22}$ .

Substituting the values in equations (6) and (7), yields

$$[\sigma(\tilde{\mu}_\alpha), \sigma(\tilde{\mu}_\delta)] = (3.96, 4.00) \text{ mas yr}^{-1}. \quad (12)$$

So, just as the error in the difference of measurements is not substantially bigger than the separate errors, so the error in the combined measurement is not substantially smaller. Again this is because of the strong correlation between the SDSS/USNO-A and USNO-B proper-motion measurements.

The upper panel of Figure 1 shows the mean  $\chi_{\text{ext}}^2$  as a function of  $\chi_{\text{int}}^2$ . In the inner core ( $\chi_{\text{int}}^2 < 4$ ), which accounts for about 70% of the sample,  $\langle \chi_{\text{ext}}^2 \rangle \sim 2$ , i.e., its expected value. It grows beyond that, as one would expect, because if  $\chi_{\text{int}}^2$  is high, then at least one of the measurements must be seriously in error, which should then corrupt the combined estimate. However, the rise unexpectedly seems to saturate at  $\langle \chi_{\text{ext}}^2 \rangle \sim 4$ , so that the most discrepant measurements in SDSS/USNO-A must be significantly anti-correlated with the the most discrepant ones in USNO-B. Further insight is gained by considering the behavior of the external errors of each catalog measurement separately (still plotted as a function of the combined internal error). These show similar behavior, indicating that they are about equally responsible for the problem measurements. They also track the combined external error in the Gaussian part of the distribution ( $\chi_{\text{int}}^2 \lesssim 8$ ), but at higher  $\chi_{\text{int}}^2$  the individual external errors continue to rise roughly linearly, while the external errors flatten. The deviation is pronounced only for  $\chi_{\text{int}}^2 \gtrsim 16$ , which encompasses only about 35 quasars. It will be interesting to see if this pattern is maintained when larger areas of the SDSS catalog are made available.

## 4. Stellar Proper Motion Catalog

### 4.1. Construction

It is now straightforward to construct the catalog. We begin with our full sample of  $5 \times 10^6$  stars selected from the SDSS catalog that satisfy  $\mu \geq 7 \text{ mas yr}^{-1}$  and  $r' < 22$ , but excluding faint ( $r' > 20$ ), high proper-motion ( $\mu > 420 \text{ mas yr}^{-1}$ ) stars for reasons described below. We first correct these proper motions for bias in the same way that we did for the quasars. See equation (1). However, since the quasars require spectroscopic confirmation, and the SDSS spectrographic catalog covers a smaller section of the sky than the photometric catalog, there are many stars that are too far from any quasars to be reliably de-biased. Therefore, we accept stars only if  $w_{\text{tot}}^m > 4$ . At this point, having obtained a better estimate of the proper motion, we further restrict the proper motion criterion to  $\mu > 10 \text{ mas yr}^{-1}$ .



These two selection procedures, primarily the first, reduce our sample by more than half to  $1.9 \times 10^6$ .

We then search for the remaining stars in USNO-B, using a  $2''$  search radius to allow for both catalog errors and the motion of the star between the SDSS epoch and the (2000) epoch of USNO-B. We increase this radius to  $2''.5$  for stars with  $\mu > 420 \text{ mas yr}^{-1}$ . As we did for the quasars, we exclude all entries with zero errors in both components, with errors greater than  $10 \text{ mas yr}^{-1}$  in either component, or with  $w_{\text{tot}}^m < 3$ . This reduces the catalog to just under  $9 \times 10^5$ .

Finally, we impose two additional cuts

$$r' \leq 20, \quad \mu > 20 \text{ mas yr}^{-1} \quad (13)$$

As we describe in § 5, the great majority of detections with  $r > 20$  are spurious. This is the reason for the first cut. The proper motion threshold is set by demanding “ $5\sigma$ ” detections. If the statistics were Gaussian, this would imply a probability  $\exp(-25/2) \sim 10^{-5.4}$  that a zero-proper-motion star could scatter into our sample. However, as shown by Figure 1, the tail of the error distribution is decidedly non-Gaussian, with roughly 1% of the quasars showing internal errors,  $\chi_{\text{int}}^2 > 16$ . While the initial evidence is that this large tail of internal errors is not reflected in a similar tail of external errors (see § 3.2 and Fig. 1), we nevertheless proceed conservatively and set a relatively high threshold to minimize contamination.

The sky coverage of the catalog is shown in Figure 2. The lightly shaded areas indicate the coverage of the photometrically selected stars, while the darker areas are the regions of the catalog. The difference is mostly due to the smaller region covered by the spectroscopic catalog, but is also affected by the failure to recover most quasars in USNO-B.

## 5. Catalog Characterization

In order to gain insight into the properties of the catalog as a function of magnitude and proper motion, we compare in Figure 3 the internal error distribution of the quasars (*open circles*) to that of various subsets of the stars (*filled circles*). For example, the lowest track compares the quasars to the stars with  $\mu \geq 20 \text{ mas yr}^{-1}$  and  $r' < 18$ . The two distributions are essentially the same, which corroborates one’s expectation that the errors are the same and that there are not a large number of spurious entries in this subsample. This is to be compared with the middle track, which shows  $\mu \geq 20 \text{ mas yr}^{-1}$  and  $19.5 < r' < 20$ . This sample has about an order of magnitude increase in the fraction of stars with large internal errors,  $\chi_{\text{int}}^2 \gtrsim 20$ . Hence, either there are a large number of spurious entries in this subset or the faint magnitudes induce unusually large proper-motion errors in one or both

catalogs. Comparison of the first and fifth tracks shows that the problem cannot be entirely faint magnitudes. These tracks represent the same magnitude range  $19 < r < 19.5$ , but very different proper-motion cuts,  $\mu > 20 \text{ mas yr}^{-1}$  and  $\mu > 100 \text{ mas yr}^{-1}$ , respectively. The higher proper-motion sample has a much higher rate of large internal errors. Thus, it is likely that spurious detections play a major role at faint magnitudes, and more so when the proper motion is also high.

It is well-known that both SDSS and USNO-B separately contain large numbers of spurious high proper-motion stars. For USNO-B, roughly 99% of the entries with  $\mu > 180 \text{ mas yr}^{-1}$  are spurious (Gould 2003c). Since SDSS is merely matching to the nearest USNO-A entry, spurious entries are bound to occur, especially when the star is too faint to be recovered by USNO-A. While the chance that spurious entries will coincide in both catalogs within our  $2''$  search radius is small in each individual case, there are literally of order  $10^6$  trials.

The implication is that one must carefully select from the catalog in accordance with one’s goals and guided by Figure 3. If for example, one wants a very clean sample, one might accept only stars in each magnitude range up to the maximum  $\chi_{\text{int}}^2$  to which the stars track the quasars. On the other hand, if one were looking for common proper motion companions to brighter stars, one might search up to  $r' \leq 20$ , recognizing that additional confirmation would be required for any candidates thus detected.

Figure 4 contains six reduced proper motion (RPM) diagrams. These have  $\mu \geq 20 \text{ mas yr}^{-1}$  and  $\mu \geq 50 \text{ mas yr}^{-1}$  thresholds for the left and right sides, respectively. All diagrams are restricted to  $r' < 19$ . The ordinates in these diagrams are the adjusted RPMs, e.g.,

$$\eta_{r'} \equiv r' + 5 \log \frac{\mu}{1'' \text{ yr}^{-1}} - 1.47 |\sin b|, \quad (14)$$

where  $b$  is Galactic latitude. The first two terms are the standard RPM. These are related to the absolute magnitude by,

$$H_{r'} = r' + 5 \log \frac{\mu}{1'' \text{ yr}^{-1}} = M_{r'} + \frac{v_{\perp}}{47.4 \text{ km s}^{-1}} + A_{r'}, \quad (15)$$

which causes an RPM diagram to simulate a color-magnitude diagram. Here  $v_{\perp}$  is the transverse velocity and  $A_{r'}$  is the extinction. However, Salim & Gould (2003) showed that while RPM diagrams at different  $b$  look similar, they are vertically offset from one another. They therefore introduced  $\eta$ , which has a  $b$ -dependent term designed to compensate for this effect.

The figure illustrates the changing morphology of RPM diagrams with choice of color and proper-motion threshold. In the “high” threshold sample, the white dwarfs (lower left),

red dwarfs (lower right), and subdwarfs (lower diagonal track) are all more pronounced. We have placed “high” in quotation marks because this sample has a lower proper-motion threshold than any previous RPM diagram at these magnitudes, except that of Digby et al. (2003) to which it is comparable. See in particular their figures 6 and 7. The low-threshold sample has a higher density of upper-main-sequence and turnoff stars. (The two sets of diagrams are drawn respectively from 9% and 60% of the catalog, so that they have a similar number of points, roughly 25,000.) Note also that in  $u' - i'$ , the red subdwarfs tend to merge with the main-sequence red dwarfs, while in  $r' - i'$ , the red ends of these tracks remain distinct.

## 6. Catalog Description

The first four columns of the catalog give the right ascension and declination of the star as reported in SDSS and USNO-B, respectively. Both are equinox 2000. USNO-B is also epoch 2000, while SDSS is at the epoch of the observations (which are circa 2000). Columns (5) and (6) give the east and north components of  $\vec{\mu}_{\text{cor},S}$ , the corrected SDSS proper motion as described by equations (1) and (4). Columns (7) and (8) give  $\vec{\mu}_{\text{cor},B}$ , the corrected USNO-B proper motion, while columns (9) and (10) give the weighted average of the two as described by equation (A5). Column (11) is the scalar proper motion, i.e., the root-sum-square of columns (9) and (10). Columns (12) through (16) give SDSS PSF photometry in  $u'$ ,  $g'$ ,  $r'$ ,  $i'$ , and  $z'$ , while columns (17) and (18) give the blue and red magnitudes of the USNO-A star that SDSS has associated with this entry. Column (19) gives  $H_{r'}$ , the unadjusted  $r'$  band RPM (see eq. [15]), while column (20) gives the adjusted RPM,  $\eta_{r'}$  (see eq. [14]). Columns (21) and (22) give  $w_{\text{tot},S}$  and  $w_{\text{tot},B}$ , the total statistical weights of the quasars used in the local proper-motion correction. See equation (1). Finally, column (23) gives the internal-error indicator  $\langle \chi_{\text{int}}^2 \rangle$  as described by equation (A3).

The fortran format statement for the catalog is (4f10.5,7f7.1,5f7.3,2f5.1,2f7.3,2f5.1,f7.2). The catalog can be found at <http://www-astronomy.mps.ohio-state.edu/~gould/SDSS-USNOB/cat.dat.gz>

If the catalog is restricted to regions in which the stars track the quasars in Figure 3, then the errors can be expected to be similar to those of the quasars, namely  $\sigma_{\mu} = 4.0 \text{ mas yr}^{-1}$  in each direction.

We thank D. Monet and the USNO-B team for providing us with a copy of the USNO-B1.0 catalog. Support for the SDSS Archive comes from the Alfred P. Sloan Foundation, the Participating Institutions, NASA, the NSF, the DoE, the Japanese Monbukagakusho, and

the Max Planck Society. AG was supported by grant AST 02-01266 from the NSF and by JPL contract 1226901. JAK was supported by an OSU Dean’s Distinguished Fellowship.

### A. Covariant forms of $\chi_{\text{int}}^2$ and $\chi_{\text{ext}}^2$

In § 3.2, we evaluated  $\chi_{\text{int}}^2$  and  $\chi_{\text{ext}}^2$  under the simplifying assumption (almost realized in practice) that the proper motion measurements in the  $\alpha$  and  $\delta$  directions are not correlated. Here we present a more general treatment. The reader is referred to Gould & An (2002) or Gould (2003d) for the mathematical underpinnings of this appendix.

We first introduce the two-vector  $A^k$ , defined by,

$$A^k \equiv \kappa_i^k a_i = (a_1 - a_2, a_3 - a_4), \quad (\text{A1})$$

where the  $2 \times 4$  constraint matrix  $\kappa_i^k$  is defined by  $\kappa_i^1 = (1, -1, 0, 0)$  and  $\kappa_i^2 = (0, 0, 1, -1)$ , and where we have introduced the Einstein summation convention. It is straightforward to show that the covariances of the  $A^k$  are given by

$$\text{cov}(A^k, A^l) = C^{kl}, \quad C^{kl} \equiv \kappa_i^k c_{ij} \kappa_j^l. \quad (\text{A2})$$

Hence, the internal errors are characterized by,

$$\chi_{\text{int},m}^2 = A_m^l B^{kl} A_m^l, \quad B \equiv C^{-1}. \quad (\text{A3})$$

Explicit evaluation of  $C^{kl}$  using equations (6) and (7),

$$\sqrt{C^{kk}} = (4.49, 4.40) \text{ mas yr}^{-1}, \quad \frac{C^{12}}{\sqrt{C^{11}C^{22}}} = -0.01. \quad (\text{A4})$$

shows that it is almost exactly the same as given in the naive analysis of § 3.2.

Next we evaluate the external errors. The combined measurement of the proper motion is found by imposing the constraint  $\kappa_i^k a_i^m \rightarrow 0$ , which yields best estimates for the parameters,  $\tilde{a}_i^m$ ,

$$\tilde{a}_i^m = a_i^m - D_m^k \alpha_i^k, \quad D_m^k \equiv B^{kl} A_m^l, \quad \alpha_i^k \equiv c_{ij} \kappa_j^k \quad (\text{A5})$$

and covariances,

$$\tilde{c}_{ij} = c_{ij} - \alpha_i^k B^{kl} \alpha_j^l. \quad (\text{A6})$$

There are only two independent components of  $\tilde{a}_i$  (one for each direction), and so  $\tilde{c}_{ij}$  is effectively a  $2 \times 2$  matrix. We therefore compress their representations,

$$\tilde{a}_{2i-1} \rightarrow \tilde{a}_i, \quad \tilde{c}_{2i-1,2j-1} \rightarrow \tilde{c}_{ij}, \quad (\text{A7})$$

and evaluate

$$\sqrt{\tilde{c}_{ii}} = (3.96, 4.00) \text{ mas yr}^{-1}, \quad \frac{\tilde{c}_{12}}{\sqrt{\tilde{c}_{11}\tilde{c}_{22}}} = -0.063. \quad (\text{A8})$$

Again, this is almost identical to the result obtained in the naive treatment. The external error is then characterized by,

$$\chi_{\text{ext},m}^2 = \sum_{i,j=1}^2 \tilde{a}_i^m \tilde{b}_{ij} \tilde{a}_j^m, \quad \tilde{b} \equiv \tilde{c}^{-1}. \quad (\text{A9})$$

## REFERENCES

- Alcock, C. et al. 2001, *ApJ*, 562, 337
- Chaname, J., & Gould, A. 2004, *ApJ* in press, astro-ph/0307434
- Digby, A.P., Hambly, N.C., Cooke, J.A., Reid, I.N., & Cannon, R.D. 2003, *MNRAS*, 344, 583
- European Space Agency (ESA). 1997, *The Hipparcos and Tycho Catalogues* (SP-1200; Noordwijk: ESA)
- Gizis, J. E. & Reid, I. N. 1997, *PASP*, 109, 849.
- Goldman, B. and The EROS Collaboration 1999, *A&A*, 351, L5.
- Gould, A. 2003 *ApJ*, 583, 765
- Gould, A. 2003, *ApJ*, 592
- Gould, A. 2003, *AJ*, 126, 472
- Gould, A. 2003, preprint, astro-ph 0310577
- Gould, A. & An, J. 2003, *ApJ*, 565, 1381
- Gould, A. & Morgan, C.W. 2003, *ApJ*, 585
- Gould, A., & Salim, S., 2003, *ApJ*, 582, 1001
- Hambly, N. C. et al. 2001, *MNRAS*, 326, 1279
- Hambly, N.C., Davenhall, A.C., Irwin, M.J., & MacGillivray, H.T. 2001, *MNRAS*, 326, 1315
- Henry, T.J., Ianna, P.A., Kirkpatrick, J.D., & Jahreiss, H. 1997, *AJ*, 114, 388
- Høg, E. et al. 2000, *A&A*, 355, L27
- Jahreiss, H., Scholz, R., Meusinger, H., & Lehmann, I. 2001, *A&A*, 370, 967
- Jones, E. M. 1972, *ApJ*, 177, 245
- Lépine, S., Shara, M.M., & Rich, R.M. 2003, *AJ*, 126, 921
- Liebert, J., Dahn, C.C., Gresham, M., & Strittmatter, P.A. 1979, *ApJ*, 233, 226
- Liebert, J., Dahn, C.C., & Monet, D.G. 1988, *ApJ*, 332, 891

- Luyten, W. J. 1979, 1980, New Luyten Catalogue of Stars with Proper Motions Larger than Two Tenths of an Arcsecond (Minneapolis: University of Minnesota Press)
- Luyten, W. J. & Hughes, H. S. 1980, Proper Motion Survey with the Forty-Eight Inch Schmidt Telescope. LV. First Supplement to the NLTT Catalogue (Minneapolis: University of Minnesota)
- Luyten, W. J. 1970, 1977, Luyten's White Dwarf Catalogues (Univ. of Minnesota, Minneapolis)
- Luyten, W. J. 1979, LHS Catalogue: a catalogue of stars with proper motions exceeding  $0''.5$  annually (Minneapolis: University of Minnesota Press)
- Monet, D. G. 1998, American Astronomical Society Meeting, 193, 112003
- Monet et al. 2003, AJ, 125, 984
- Oppenheimer, B.R., Hambly, N.C., Digby, A.P., Hodgkin, S.T., & Saumon, D. 2001, Science, 292, 698
- Reid, I.N. & Cruz, K.L. 2002, AJ, 123, 2806
- Reid, I. N., Liebert, J., & Schmidt, G.D. 2001, ApJ, 550, L61
- Reid, I.N., Sahu, K.C., & Hawley, S.L. 2001, ApJ, 559, 942
- Ruiz, M. T., Wischnjewsky, M., Rojo, P.M., & Gonzalez, L.E., 2001 ApJS, 133, 119
- Ryan, S.G. 1992, AJ, 104, 1144
- Salim, S. & Gould, A. 2000, ApJ, 539, 241
- Salim, S., & Gould, A. 2003 ApJ, 582, 1011
- Schmidt, G. D., Liebert, J., Harris, H.C., Dahn, C.C., & Leggett, S.K. 1999, ApJ, 512, 916
- Scholz, R.-D., Meusinger, H., & Jahreiss, H. 2001, A&A, 374, L12
- Sumi, T. et al., 2003, astro-ph/0305315
- Wroblewski, H. & Costa, E., 2001, A&A, 367, 725
- Yoo, J., Chaname, J., Gould, A., 2004, ApJ, in press, astro-ph/0307437
- Zacharias, N. et al. 2000, AJ, 120, 2131.

Zacharias, N., Zacharias, M.I., Urban, S.E., Rafferty, T.J.,. 2001, AAS, meeting #129.08,  
vol 33, p1495



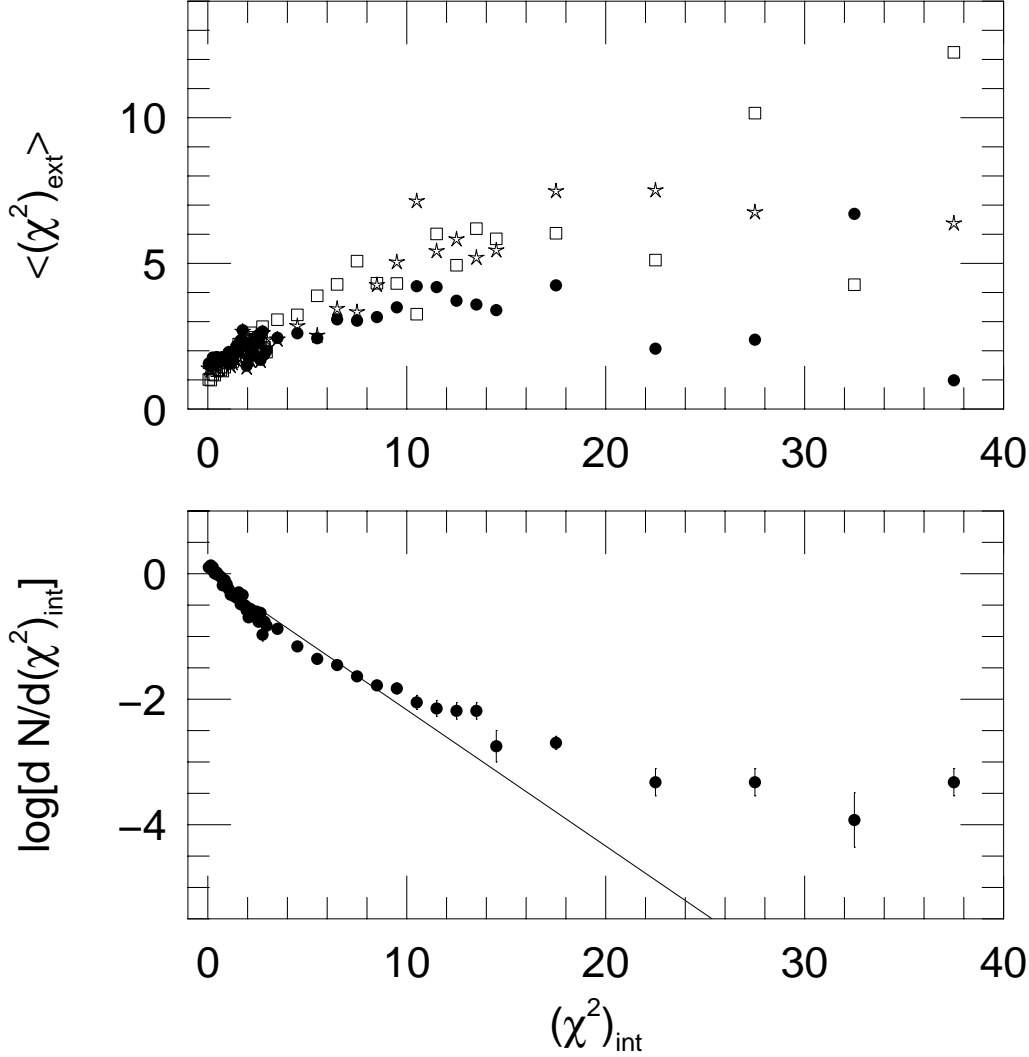


Fig. 1.— Lower panel: Distribution of the internal-error indicator  $\chi_{\text{int}}^2$  (derived from the difference between SDSS and USNO-B proper-motion measurements, see eq. [A3]) for  $N = 3368$  quasars. The distribution is normalized by  $N/2$  to permit easy comparison with other distributions (see Fig. 3). The straight line is the expected distribution assuming Gaussian statistics. It is not a fit: it has no free parameters. The distribution is Gaussian in its core ( $\chi_{\text{int}}^2 \lesssim 8$ ) but deteriorates significantly in its wings. Upper panel: Mean external-error indicator  $\chi_{\text{ext}}^2$  (derived from the differences between the measured proper motions and their known zero value, see eq. [A9]) as a function of  $\chi_{\text{int}}^2$  for the same sample. Values for the SDSS-only (*stars*), USNO-B-only (*open squares*) and combined (*filled circles*) measurements are shown separately. In the inner Gaussian core  $\chi_{\text{int}}^2 \lesssim 4$ , all three cluster around their expected values of 1, 1, and 2 respectively. At higher  $\chi_{\text{int}}^2$ , for which SDSS and USNO-B are in increasing conflict and therefore cannot both be correct,  $\chi_{\text{ext}}^2$  rises as would be expected. Surprising, however, for the combined measurements, the rise saturates at  $\chi_{\text{ext}}^2 \sim 4$ , perhaps indicating that large-error measurements of SDSS and USNO-B are anti-correlated.

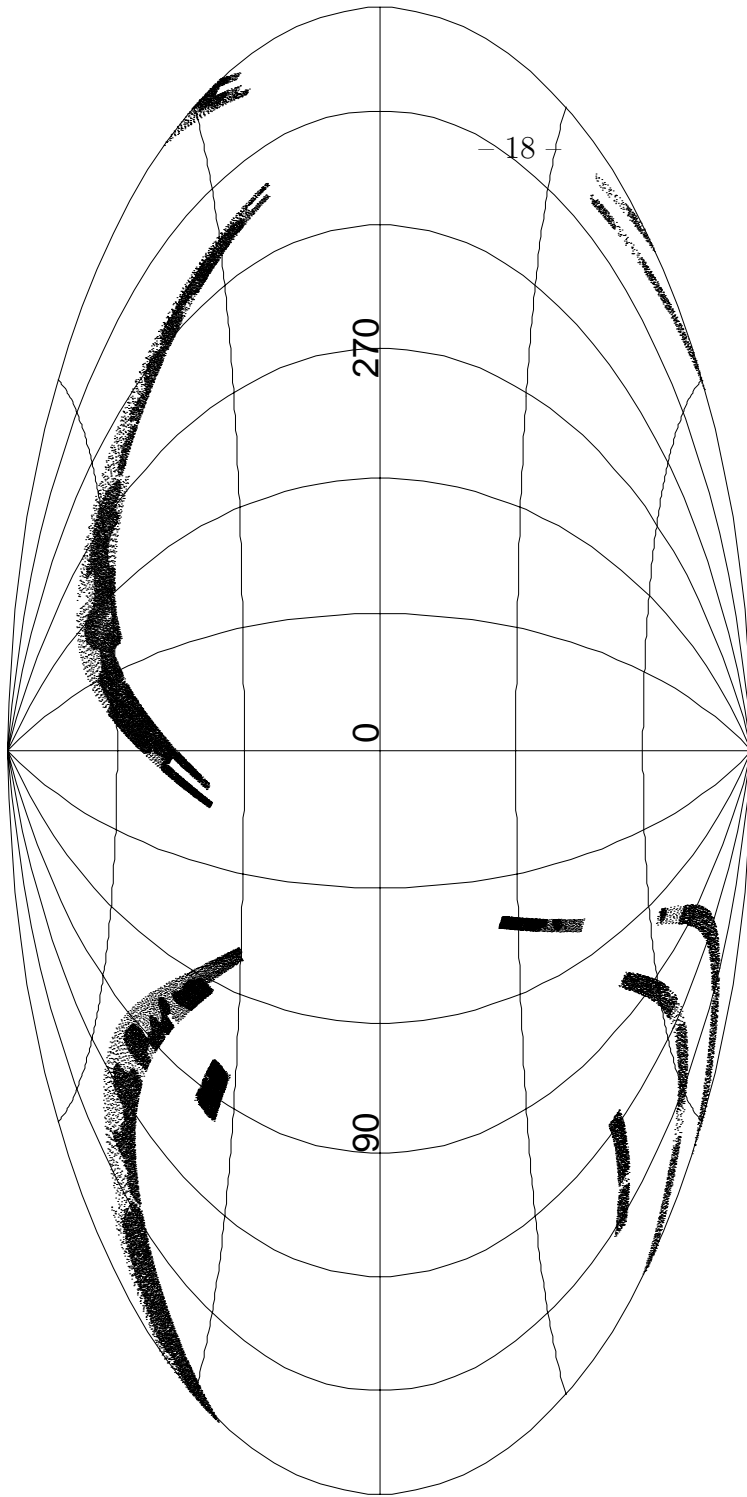


Fig. 2.— Distribution of catalog stars on the sky. The lightly shaded regions are the original selection from the SDSS photometric database. The heavily shaded regions are the final sample. Their smaller area reflects the fact that the proper motions must be corrected using SDSS quasars in the spectroscopic database, which has substantially smaller sky coverage.

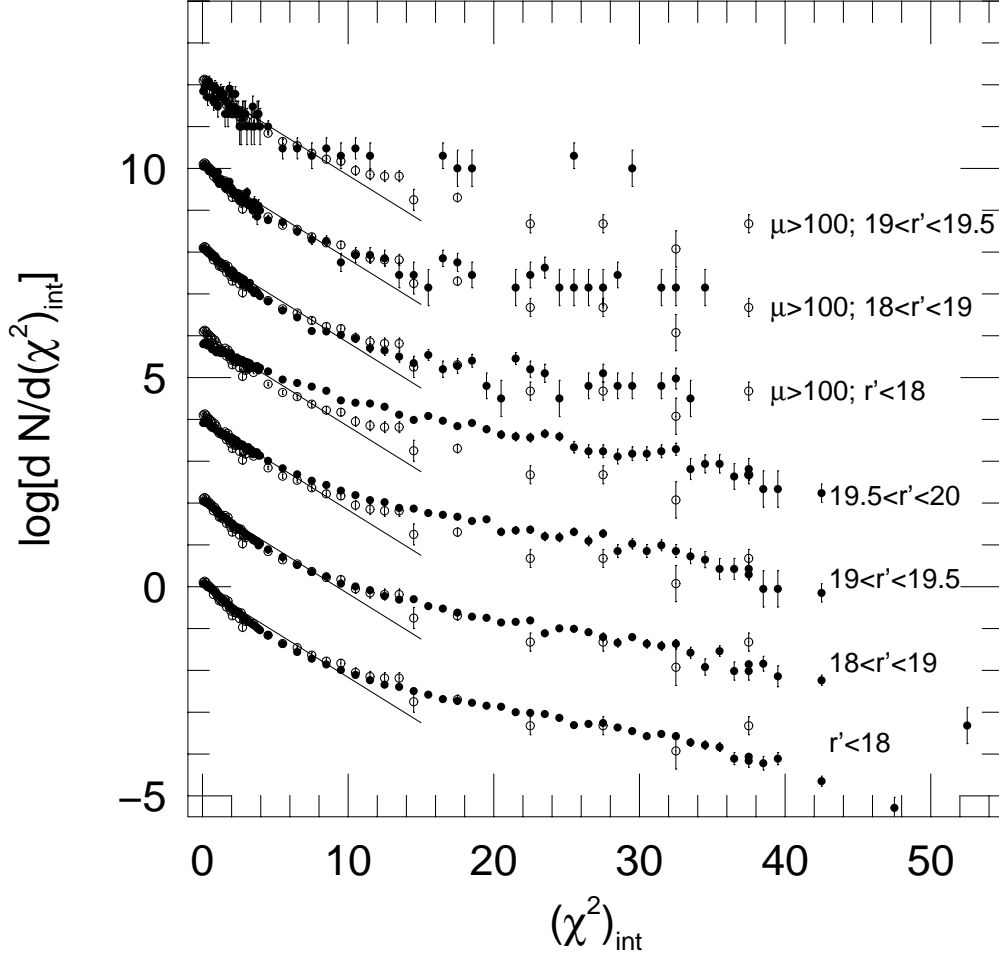


Fig. 3.— Distribution of internal-error indicator  $\chi_{\text{int}}^2$  for various subsets of the stellar catalog (*solid circles*) compared to that for quasars (*open circles*, see Fig. 1). The three upper tracks are restricted to  $\mu > 100 \text{ mas yr}^{-1}$ , while the four lower tracks go down to the catalog limit of  $\mu = 20 \text{ mas yr}^{-1}$ . Note that the  $r' < 18$  distributions track the quasars extremely well, indicating that their error distributions are similar and that there are few spurious objects. However, the faintest bin,  $19.5 < r' < 20$ , has a tail that is an order of magnitude higher than the quasars, probably indicating a large number of spurious objects, and this effect is even more pronounced among the high proper-motion stars.

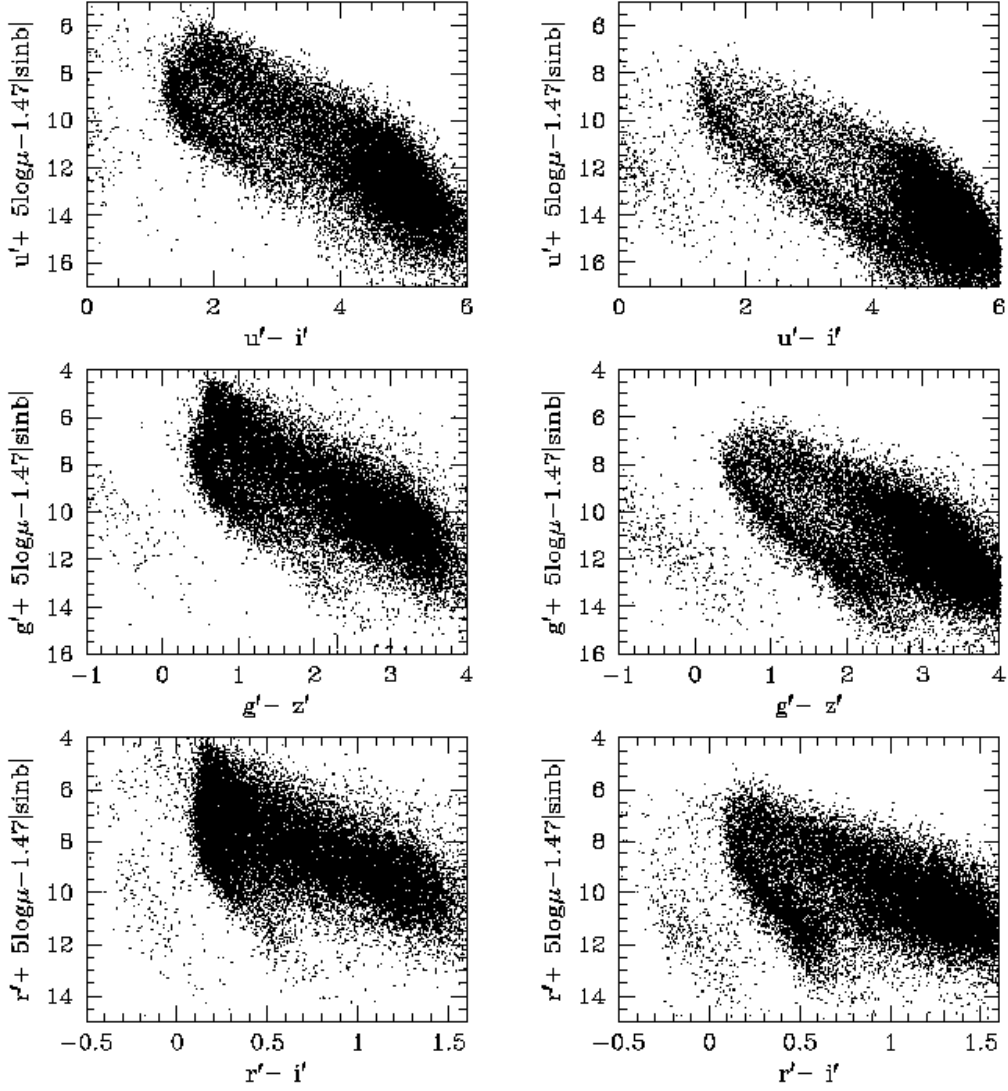


Fig. 4.— Reduced proper-motion diagrams for catalog stars  $r' < 19$  for stars down to the proper-motion limit of  $20 \text{ mas yr}^{-1}$  (left panels) and restricted to relatively high proper motions,  $\mu \geq 5 \text{ mas yr}^{-1}$  (right panels). All panels have about 25,000 points, meaning that the right panels are drawn from a larger fraction of the catalog. The right panels are relatively enhanced in white dwarfs (lower left), red dwarfs (lower right), and subdwarfs (lower diagonal track). Note also how the subdwarf morphology evolves with color.